

Massive statistical analysis of autoscaled data: the case of the double reflection signature in mid-latitude vertical ionograms

M. Pezzopane and C. Scotto

Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143, Rome, Italy

Correspondence to: M. Pezzopane (michael.pezzopane@ingv.it)

Abstract

This work shows how new capabilities can emerge from a massive statistical analysis of previously overlooked autoscaled data. In particular, the paper shows how autoscaling methods for vertical ionograms, specifically Autoscala, can offer a new kind of data that are not currently available at World Data Center or elsewhere and not reported by manual ionogram scalars. In this context, an example of such new analyses is the presentation of a statistics of occurrence of the double reflection phenomenon that sometimes characterizes ionograms. In order to establish this original statistics, a method developed to smooth out a specific autoscaling problem was utilized, and a large data set of ionograms recorded from 2003 to 2008 by the AIS-INGV ionosondes installed at the ionospheric stations of Rome (41.8°N, 12.5°E) and Gibilmanna (37.9°N, 14.0°E), Italy, was analyzed. The main results that emerged from the study are hence illustrated and briefly discussed.

Keywords: autoscaling system; statistical analysis; ionogram; Autoscala; mid-latitude ionosphere

1. Introduction

Rapid growth of sensor instrumentation networks has created a unique challenge for the modern scientist, coined in the literature as “Big Data” (Lynch, 2008). While it is technically feasible to store *zettabyte* volumes of the Big Data, their analysis has become impractical without intelligent software tools that mimic human capability to discover and interpret data patterns. Several domains of physics have already faced the challenges of massive sensor data acquisition and analysis, starting as early as the 1950s; the experience collected in dealing with the challenges of Big Data Analytics in these domains is directly applicable to the modern world of science.

In particular, ground-based observation of the Earth's ionosphere by means of the high-frequency (HF) vertical sounding with an ionosonde (Davies, 1990) is one of the longest standing observational techniques known for its large imagery capacity and associated intelligent software system development (Wilkinson, 1998) that can be traced back into the early 1960s. The HF ionosonde transmits radio wave signals vertically towards the ionosphere and records the travel time of the observed echoes that is then used to calculate distance to the point of signal reflection in the ionosphere. The main visualization product of the ionosonde observations is ionogram, a display of the echo strength as function of the signal carrier frequency, commonly varied from 1 to 20 MHz, and the signal travel time presented as altitude in km.

Analysis of ionograms for signatures of ionospheric radio wave propagation continues to be one of the most accurate and effective ways to monitor and study plasma dynamics of fundamental importance to Space Weather research and applications. Detection and interpretation of the ionogram signatures is known in the literature as “scaling”; immediate near-real-time availability of well-scaled ionogram-derived data has become crucial in recent decades. Since the 1960s multiple groups of researchers have addressed the problem referred to in the ionospheric community as the “ionogram autoscaling”, and appropriate algorithms have been developed that can automatically scale the echo trace recorded by an ionosonde and output the standard ionospheric characteristics (Reinisch and Huang, 1983; Fox and Blundell, 1989; Igi et al., 1993; Tsai and Berkey, 2000; Zabotin et al., 2006; Ding et al., 2007; Su et al., 2012).

The human brain has however demonstrated an as yet unsurpassed capacity to appropriately group dispersed echoes of the same type, to recognize fading and evanescent traces, and to extrapolate, where necessary, missing echoes to obtain the values of the ionospheric characteristics. Therefore, manual data obtained by experienced ionogram scalers are still considered the most reliable for research in the morphology and physics of the ionosphere. Several studies compared autoscaled and manually scaled values, demonstrating that the challenges arising in this kind of intelligent system development are varied and complex (Gilbert and Smith, 1988; Jacobs et al., 2004; McNamara, 2006; Pezzopane and Scotto, 2005, 2007; Bamford et al., 2008; Stanimir et al., 2012). Continuing development of better autoscaling algorithms steadily improves the reliability and accuracy of the automatically scaled data (Reinisch et al., 2005; Galkin and Reinisch, 2008; Scotto and Pezzopane, 2008; Pezzopane and Scotto, 2010).

In this work, we will depart from the classic approach of a “yet better autoscaling algorithm” to explore the topic that has been in the spotlight of intelligent systems community: how smart software can be instrumental in discovering new knowledge. We will present an example how automatically derived data, with their inherent uncertainties of interpretation, can aid studies in ionospheric morphology by employing statistical analysis and looking at characteristics far from the mainstream of human attention. We will base our analysis on a byproduct of ionogram processing by the anti double reflection filter developed by Scotto and Pezzopane (2008) to establish a viable statistics of occurrence of “second hops” of the signal between the ionosphere and the Earth. While little attention is paid to this feature in manual interpretation of ionograms, we will employ the autoscaling double reflection filter to analyze a large data set of mid-latitude ionograms recorded at Rome (41.8° N, 12.5° E) and Gibilmanna (37.9° N, 14.0° E), Italy, by the advanced ionospheric sounder designed and developed at the Istituto Nazionale di Geofisica e Vulcanologia (AIS-INGV) (Zuccheretti et al., 2003).

2. The double reflection traces and the software Autoscala

Ionogram records are inspected for presence of the echo traces corresponding to reflection of the sensor signal from the layers of ionospheric plasma (Fig. 1); at times, ionograms also exhibit a family of similarly shaped traces appearing at double, triple, etc. altitudes. The higher order reflections do not indicate physical presence of plasma structures at integer multiples of the altitude; they are due to “multi-hops” of the signal that travels between the ionosphere and the Earth more than once (Fig. 1). The multi-hop reflections from the F region at night were first investigated by Pierce and Mimno (1940) who concluded that timeline patterns of varying multiplicity of the hops can be associated with dynamics of particular plasma curvatures in the F region causing focusing effects that amplified the signal to support higher orders of its propagation. Bowman (1964) later concentrated on cases of high-order multiple reflections (over 10 hops), arguing that geometry of the ionospheric irregularities present when the phenomenon was observed ruled out the possibility of focusing effects. Instead, he presented evidence in support of a mechanism involving the decrease of non-deviative ionospheric absorption to sustain high order multi-hop propagation.

When an ionogram shows multiple reflections, autoscaling systems like Autoscala or ARTIST can be misled, especially when curved isodensity contours add focusing effects, and the second-order reflection can be identified as first order. When this happens, even though autoscaled values like the critical frequency of the F2 layer (f_oF2) are not affected by a

significant error, autoscaled values like the propagation factor ($M(3000)F_2$) and the maximum usable frequency for a 3000 km circuit ($MUF(3000)F_2$) are strongly underestimated, and this error is reflected in the deviation of the estimated electron density profile from the true one.

Fox and Blundell (1989), Igi et al. (1993), and Tsai and Berkey (2000) discussed this subject in their works. Nevertheless, they underlined only the issue of the blanketing and overlapping of the F trace caused by E sporadic (Es) multiple traces, without considering cases in which only second-order reflections of the F trace are present. For instance, Fox and Blundell (1989) in their autoscaling system included a check test looking only for multiple Es traces overlapping the F trace. The F trace identification performed by Autoscala is not affected by Es multiple reflections, because the algorithm is designed to recognize the typical vertical asymptotical behaviour of the F trace (Pezzopane and Scotto, 2007, 2008). On the other hand, Autoscala could be misled by multiple reflections of the F trace. For this reason, Scotto and Pezzopane (2008) developed a simple filter to smooth out this problem, making Autoscala output more reliable.

The basic idea of the filter is to compare the images contained in two sliding rectangles whose points are placed on the ionogram at frequencies and virtual heights so that the filter takes into account the vertical stretching exhibited by the second hop trace (see Fig. 3a of Scotto and Pezzopane, (2008)). The two rectangles are slid all over the ionogram by varying the frequency and the virtual height of their centres. At each sliding step the images contained in the two rectangles are compared using the correlation method, and if they are found similar the image in the upper rectangle is considered a second-order reflection trace, this info being archived in a proper file, and at the end of the sliding process the second-order reflection trace is appropriately deleted from the ionogram (see Fig. 3b of Scotto and Pezzopane, (2008)).

3. Results and Discussion

The analysis was based on the ionograms recorded at Rome and Gibilmanna from 2003 to 2008 by the AIS-INGV ionosonde. For each year, this data set was split according to the northern hemisphere seasons centred on the corresponding solstices and equinoxes: Autumn from 5 August to 4 November, Winter from 5 November to 4 February, Spring from 5 February to 4 May, and Summer from 5 May to 4 August. Hence, for each season of each year, the corresponding ionograms were automatically scaled by the version of Autoscala including the anti double reflection filter.

Fig. 2 shows the percentage of occurrence of double reflections for Gibilmanna in autumn for different years in the form of histograms, as detected by the anti double reflection filter

implemented in Autoscala. The y-axis represents the percentage of occurrence, and the x-axis the solar zenith angle (χ) in intervals of five degrees. For example, a full bar corresponding to a χ equal to 60° refers to ionograms recorded when $57.5^\circ \leq \chi < 62.5^\circ$. The percentage of occurrence was calculated as the number of ionograms for which the second-order reflection was detected divided by the total number of ionograms included in that χ interval. In order to have a significant statistics, the percentage of occurrence was calculated only when the total number of ionograms was greater or equal to 20. The histograms referring to ionograms recorded from 00:00 local time (LT) to 11:59 LT are reported on the left (in grey), while on the right (in black) the histograms refer to ionograms recorded from 12:00 LT to 23:59 LT. Therefore, the morning solar terminator (Somsikov, 2011) corresponds to $\chi=90^\circ$ in the grey histograms, while the evening solar terminator to $\chi=90^\circ$ in the black histograms. Yearly smoothed sunspot number (R_{12}) is also included within each histogram.

Figs 3 and 4 report the seasonal percentage of occurrence of double reflection for Gibilmanna and Rome, for years of very low solar activity ($R_{12}(2006)=15$, $R_{12}(2007)=8$, $R_{12}(2008)=3$).

Fig. 2 shows that: a) the percentage of occurrence increases as solar activity decreases; b) the percentage of occurrence is roughly directly proportional to χ ; c) two minima, increasingly pronounced as solar activity decreases, characterize the two sets of histograms: a minimum at solar zenith angles around sunrise, and another minimum at solar zenith angles around sunset.

Figs 3 and 4 show that: d) the percentage of occurrence is practically absent for very low values of χ in spring and in summer; e) the percentage of occurrence has a maximum in winter and a minimum in summer; f) from a seasonal perspective, the minimum at sunrise is nearly constant, whereas there is a shift of sunset minimum towards higher values of χ in summer and lower values of χ in winter.

Features a), b), d), and e) are easily explainable in terms of ionization caused by solar radiation which, during daytime and high solar activity, maximizes the electron density in the lower part of the ionosphere, consequently increasing absorption of electromagnetic waves transmitted by an ionosonde.

In order to speculate about a potential physical mechanism responsible for features c) and f), these being the two minima around sunrise and sunset and their seasonal behaviour as evidenced by Figs 3 and 4, Figs 5 and 6 show, as a function of χ , the mean of the vertical electron density profiles given as output by Autoscala (the real heights corresponding to plasma frequencies equal to 3.0 MHz and 3.5 MHz are reported in light and dark green respectively) over the same periods considered in Figs 3 and 4. In Figs 5 and 6 it is seen that,

on average, for solar zenith angles before the morning solar terminator there is a rapid descent of the electron isodensity surfaces, while after the evening solar terminator there is a rapid uplift of the electron isodensity surfaces. The rapid descent and uplift of the electron isodensity surfaces occurring just before and after the solar terminators is explained in terms of the rapid production and recombination of electrons characterizing the lower F region (Rishbeth and Garriott, 1969). Observing the arrows in Figs 3, 4, 5, and 6, it is possible to note a close correspondence between the minima in the occurrence of double reflection and the rapid descents and uplifts of the electron isodensity surfaces. This correspondence can be explained considering that, for these solar zenith angles, there is an increased probability of markedly tilted electron isodensity surfaces vertically above an observer. These tilted surfaces probably backscatter the radio signal far from the transmitter and are not at all compatible with curvatures producing focusing effects, as suggested by Pierce and Mimno (1940).

Concerning the seasonal shift characterizing the sunset minimum, it may be interesting to refer to the work of Zhang et al. (1999), in which the hourly trend of the height of the electron density maximum of the F2 layer ($hmF2$) is analyzed at mid latitudes. This trend can be used to infer the probability of tilted electron isodensity surfaces above a given point. Zhang et al. (1999) discussed $hmF2$ variations as measured by the Japanese Middle and Upper atmosphere radar, and as calculated by ionosonde data - using the formula of Shimazaki (1955) - and as predicted by the International Reference Ionosphere (Bilitza and Reinisch, 2008). In their work, the authors observed that the morning drop in $hmF2$ correlates closely with χ independently of season, which explains why sunrise minimum of occurrence of double reflection does not show any seasonal shift as a function of χ , as evidenced in Figs 3 and 4. Conversely, the evening rise in $hmF2$ was shown to appear late in summer, relative to winter, and this may explain the seasonal shift as a function of χ characterizing sunset minimum of occurrence of double reflection, as seen in Figs 3 and 4.

4. Summary

Existing autoscaling software are focused on getting a reasonable imitation of how human operators scale ionograms, and they are primarily exploited to obtain a real-time monitoring of the ionosphere. This paper pointed out that autoscaling systems can also be used in a post-process mode to perform a massive statistical analysis that can give useful information about previously overlooked ionogram features. Specifically, this work demonstrated how a method developed to smooth out a specific autoscaling problem made it possible to accomplish an

original study: a statistical analysis of the occurrence of the double reflection trace that sometimes appears on an ionogram.

The most remarkable features highlighted by this analysis are the occurrence of two minima for solar zenith angles around sunrise and sunset, and a seasonal shift characterizing sunset minimum. We tried to give an explanation of both features respectively in terms of the rapid vertical movements of the electron isodensity surfaces just before and after the solar terminators, and in terms of the delayed evening uplift of the F region, passing from winter to summer. We are however aware of the fact that making a solid connection between the radio aspects of observation and the underlying atmospheric phenomena is not an easy task, that in case it needs further quantitative studies, such as for example simulation investigations, to be looked into with a sufficient detail.

Finally, it would be however interesting to verify whether the features emerging from this mid-latitude study are generalized, and it is our intention to apply this new autoscaling post-process method to perform a massive statistical analysis on ionograms recorded at high and low latitudes.

The reliability of the software for automatic interpretation of ionograms is likely to improve due to both the growing commitment of the research groups, and the constant increase in computing power. This fact, together with the capability of producing large volumes of data, and the possibility of eliminating random zero mean errors, already offers automatic data analysis considerable potential, which is bound to extend in the future.

References

- Bamford, R. A., Stamper, R., Cander, L. R., 2008. A comparison between the hourly autoscaled and manually scaled characteristics from the Chilton ionosonde from 1996 to 2004. *Radio Science* 43, RS1001, doi:10.1029/2005RS003401.
- Bilitza, D., Reinisch, B. W., 2008. International Reference Ionosphere 2007: Improvements and new parameters. *Advances in Space Research* 42(4), 599-609, doi:10.1016/j.asr.2007.07.048.
- Bowman, G. G., 1964. Ionospheric irregularities and high multiple reflections. *Australian Journal of Physics* 17, 480–489.
- Davies, K., 1990. *Ionospheric Radio*. IEE Electromagnetic Waves Series 31. London, UK: Peter Peregrinus Ltd. on behalf of the Institution of Electrical Engineers, pp 580, ISBN 0-86341-186-X.
- Ding, Z., Ning, B., Wan, W., Liu, L., 2007. Automatic scaling of F2-layer parameters from ionograms based on the empirical orthogonal function (EOF) analysis of ionospheric electron density. *Earth Planets and Space* 59(1), 51–58.
- Fox, M. W., Blundell, C., 1989. Automatic scaling of digital ionograms. *Radio Science* 24(6), 747–761.
- Galkin, I. A., Reinisch, B. W., 2008. The new ARTIST 5 for all Digisondes. *Ionosonde Network Advisory Group Bulletin* 69, pp. 1–8, IPS Radio and Space Serv., Surry Hills, N. S. W., Australia. [Available at <http://www.ips.gov.au/IPSHosted/INAG/web-69/2008/artist5-inag.pdf>.]
- Gilbert, J. D., Smith, R. W., 1988. A comparison between the automatic ionogram scaling system ARTIST and the standard manual method. *Radio Science* 23(6), 968–974.
- Igi, S., Nozaki, K., Nagayama, M., Ohtani, A., Kato, H., Igarashi, K., 1993. Automatic ionogram processing systems in Japan, in *Proceedings of Session G6 at the XXIVth General*

Assembly of the International Union of Radio Science (URSI), Kyoto, Japan, August 25 - September 2, Int. Union of Radio Sci., Ghent, Belgium.

Jacobs, L., Poole, A. W. V., McKinnel, L. A., 2004. An analysis of automatically scaled F1 layer data over Grahamstown, South Africa. *Advances in Space Research* 34(9), 1949–1952.

Lynch, C., 2008. Big Data: how do your data grow? *Nature* 455, 28-29 doi:10.1038/455028a.

McNamara, L. F., 2006. Quality figures and errors bars for autoscaled Digisonde vertical incidence ionograms. *Radio Science* 41, RS4011, doi:10.1029/2005RS003440.

Pezzopane, M., Scotto, C., 2004. Software for the automatic scaling of critical frequency foF2 and MUF(3000)F2 from ionograms applied at the Ionospheric Observatory of Gibilmanna. *Annals of Geophysics* 47(6), 1783–1790.

Pezzopane, M., Scotto, C., 2005. The INGV software for the automatic scaling of foF2 and MUF(3000)F2 from ionograms: A performance comparison with ARTIST 4.01 from Rome data. *Journal of Atmospheric and Solar-Terrestrial Physics* 67(12), 1063–1073.

Pezzopane, M., Scotto, C., 2007. The automatic scaling of critical frequency foF2 and MUF(3000)F2: A comparison between Autoscala and ARTIST 4.5 on Rome data. *Radio Science* 42, RS4003, doi:10.1029/2006RS003581.

Pezzopane, M., Scotto, C., 2008. A method for automatic scaling of F1 critical frequencies from ionograms. *Radio Science* 43, RS2S91, doi:10.1029/2007RS003723.

Pezzopane, M., Scotto, C., 2010. Highlighting the F2 trace on an ionogram to improve Autoscala performance. *Computers & Geosciences*, 36, 1168-1177, doi:10.1016/j.cageo.2010.01.010.

Pierce, J.A., Mimno, H.R., 1940. The reception of radio echoes from distant ionospheric irregularities. *Physical Review* 57, 95–105.

Reinisch, B. W., Huang, X., 1983. Automatic calculation of electron density profiles from digital ionograms: 3. Processing of bottomside ionograms. *Radio Science* 18(3), 477–492.

Reinisch, B. W., Huang, X., Galkin, I. A., Paznukhov, V., Kozlov, A., 2005. Recent advances in real-time analysis of ionograms and ionospheric drift measurements with digisondes. *Journal of Atmospheric and Solar-Terrestrial Physics* 67(12), 1054–1062.

Rishbeth, H., Garriott, O. K., 1969. Introduction to ionospheric physics. International Geophysics Series, New York, Academic Press, 331pp.

Scotto, C., Pezzopane, M., 2008. Removing multiple reflections from the F2 layer to improve Autoscala performance. *Journal of Atmospheric and Solar-Terrestrial Physics* 70(15), 1929–1934.

Shimazaki, T., 1955. World-wide daily variations in the height of the maximum electron density of the ionospheric F2-layer. *Journal of Radio Research Laboratories* 2(7), 85–97.

Somsikov, V. M., 2011. Solar terminator and dynamic phenomena in the atmosphere: A review. *Geomagnetism and Aeronomy* 51(6), 707-719, doi:10.1134/S0016793211060168.

Stanimir, M. S., Jodogne J. C., Kutiev I., Stegen K., Warnant R., 2012. Evaluation of automatic ionogram scaling for use in real-time ionospheric density profile specification: Dourbes DGS-256/ARTIST-4 performance. *Annals of Geophysics* 55(2), 283-291, doi:10.4401/ag-4976.

Su, F., Zhao, Z., Li, S., Yao, M., Chen, G., Zhou, Y., 2012. Signal Identification and Trace Extraction for the Vertical Ionogram. *Geoscience and Remote Sensing Letters* 9(6), 1031-1035, doi:10.1109/LGRS.2012.2189350.

Tsai, L. C., Berkey, F. T., 2000. Ionogram analysis using fuzzy segmentation and connectedness techniques. *Radio Science* 35(5), 1173–1186.

Wilkinson, P., Editor, 1998. Computer-Aided Processing of Ionograms and Ionosonde Records. Proceedings of Session G5 at the XXVth General Assembly of the International Union of Radio Science (URSI), Lille, France, August 28-September 5, 1996. World Data Center A for Solar-Terrestrial Physics, Report UAG-105, National Geophysical Data Center, Boulder, CO.

Zabotin, N. A., Wright, J. W., Zhabankov, G. A., 2006. NeXtYZ: three-dimensional electron density inversion for dynasonde ionograms. *Radio Science* 41, RS6S32, doi:10.1029/2005RS003352.

Zhang, S.-R., Fukoa, S., Oliver, W. L., Otsuka, Y., 1999. The height of the maximum ionospheric electron density over the MU radar. *Journal of Atmospheric and Solar-Terrestrial Physics* 61(18), 1367–1383.

Zuccheretti, E., Tutone, G., Sciacca, U., Bianchi, C., Arokiasamy, B. J., 2003. The new AIS-INGV digital ionosonde. *Annals of Geophysics* 46(4), 647–659.

Fig. 1. Ionogram recorded the 4 December 2007 by the AIS-INGV ionosonde installed at Rome, and characterized by the appearance of a second-order reflection of the F2 layer. The real trace is due to a single ionosphere-Earth hop between the transmitter (T) and the receiver (R), while the second-order trace is due to a double ionosphere-Earth hop between T and R.

Fig. 2. Percentage of occurrence of double reflection for Gibilmanna in autumn from 2003 to 2008, as a function of χ . Grey histograms refer to ionograms recorded from 00:00 LT to 11:59 LT, black histograms refer to ionograms recorded from 12:00 LT to 23:59 LT. The R_{12} yearly smoothed sunspot number is included in each histogram. Red arrows indicate the position of the minima around sunrise and sunset.

Fig. 3. Percentage of occurrence of double reflection for Gibilmanna from Winter 2007/2008 to Autumn 2008, as a function of χ . Green and red vertical lines mark the position of sunrise and sunset minima respectively.

Fig. 4. Percentage of occurrence of double reflection for Rome from Winter 2006/2007 to Autumn 2007, as a function of χ . Green and red vertical lines mark the position of sunrise and sunset minima respectively.

Fig. 5. Mean of the vertical electron density profiles, as a function of χ , autoscaled at Gibilmanna by Autoscala from Winter 2007/2008 to Autumn 2008. Real heights are shown corresponding to plasma frequency equal to 3.0 MHz (light green) and 3.5 MHz (dark green), with red arrows highlighting the rapid descents/uplifts affecting the electron isodensity surfaces.

Fig. 6. Mean of the vertical electron density profiles, as a function of χ , autoscaled at Rome by Autoscala from Winter 2006/2007 to Autumn 2007. Real heights are shown corresponding to plasma frequency equal to 3.0 MHz (light green) and 3.5 MHz (dark green), with red arrows highlighting the rapid descents/uplifts affecting the electron isodensity surfaces.

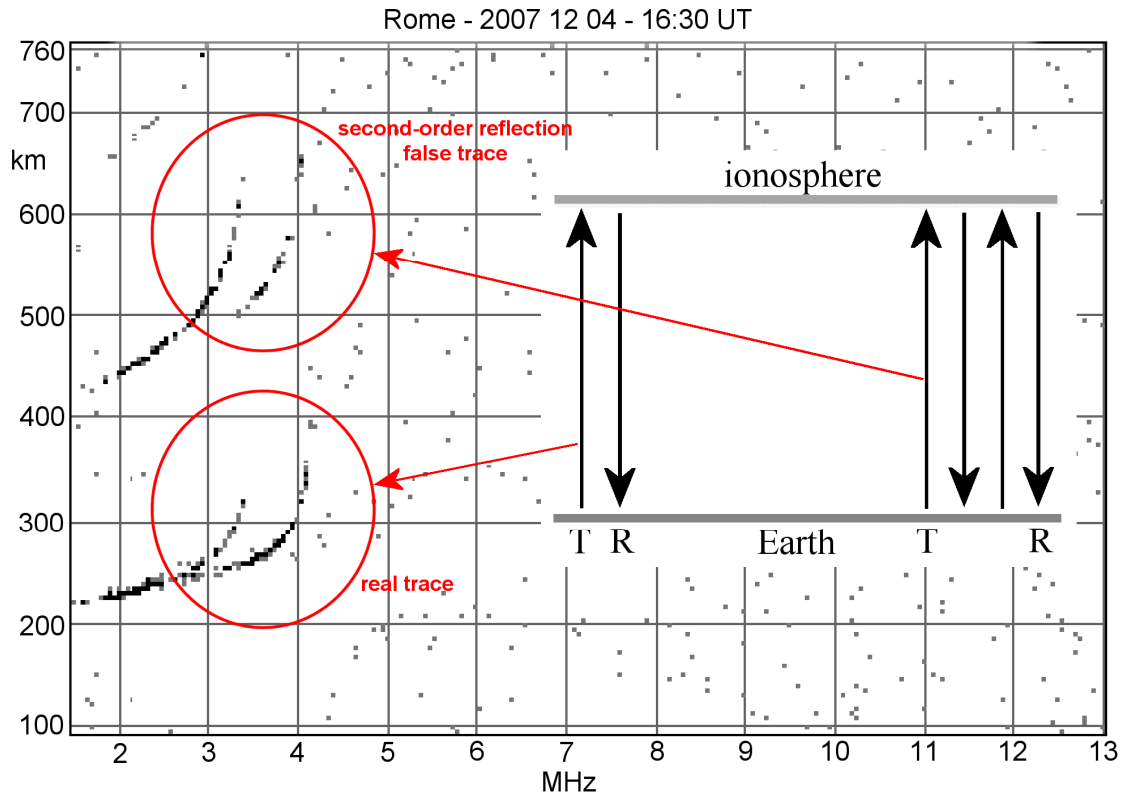


Figure 1

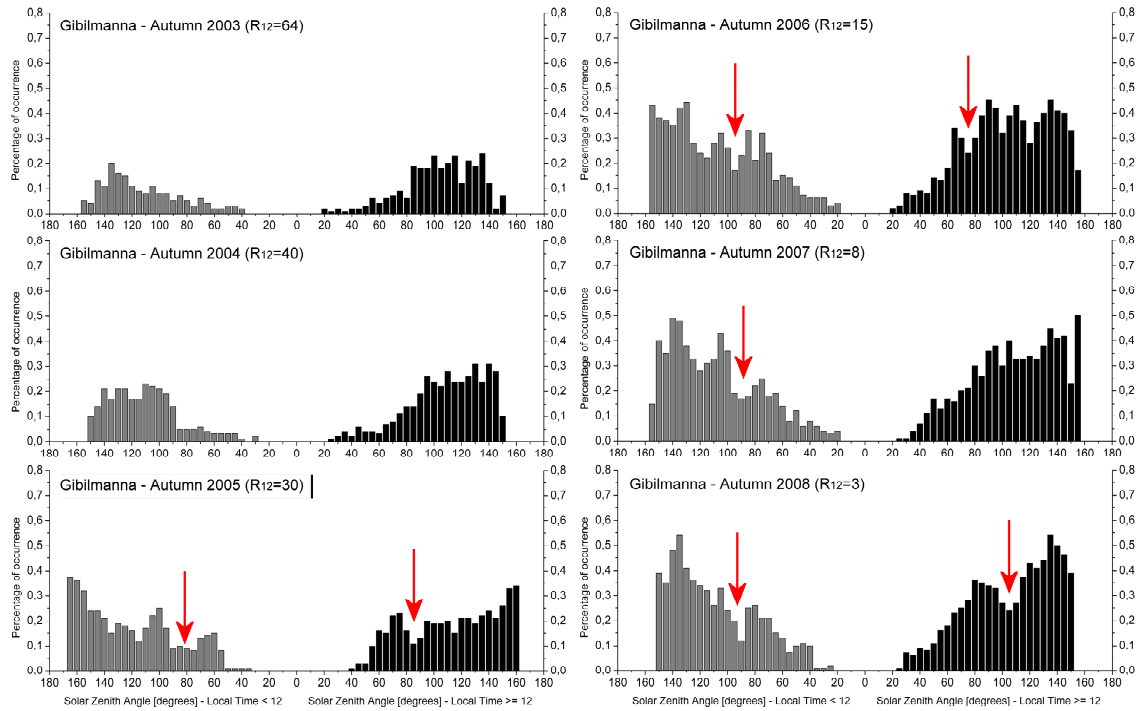


Figure 2

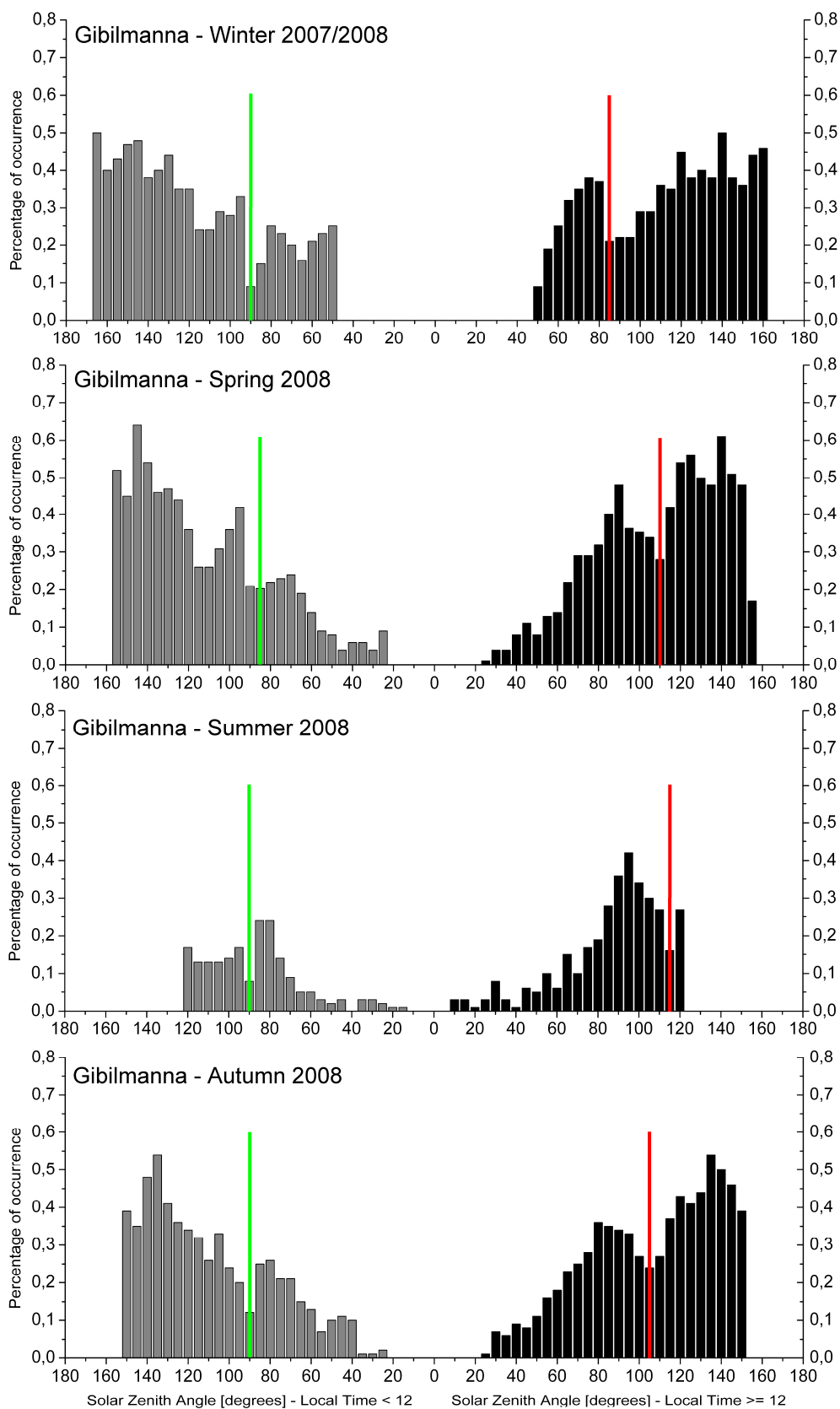


Figure 3

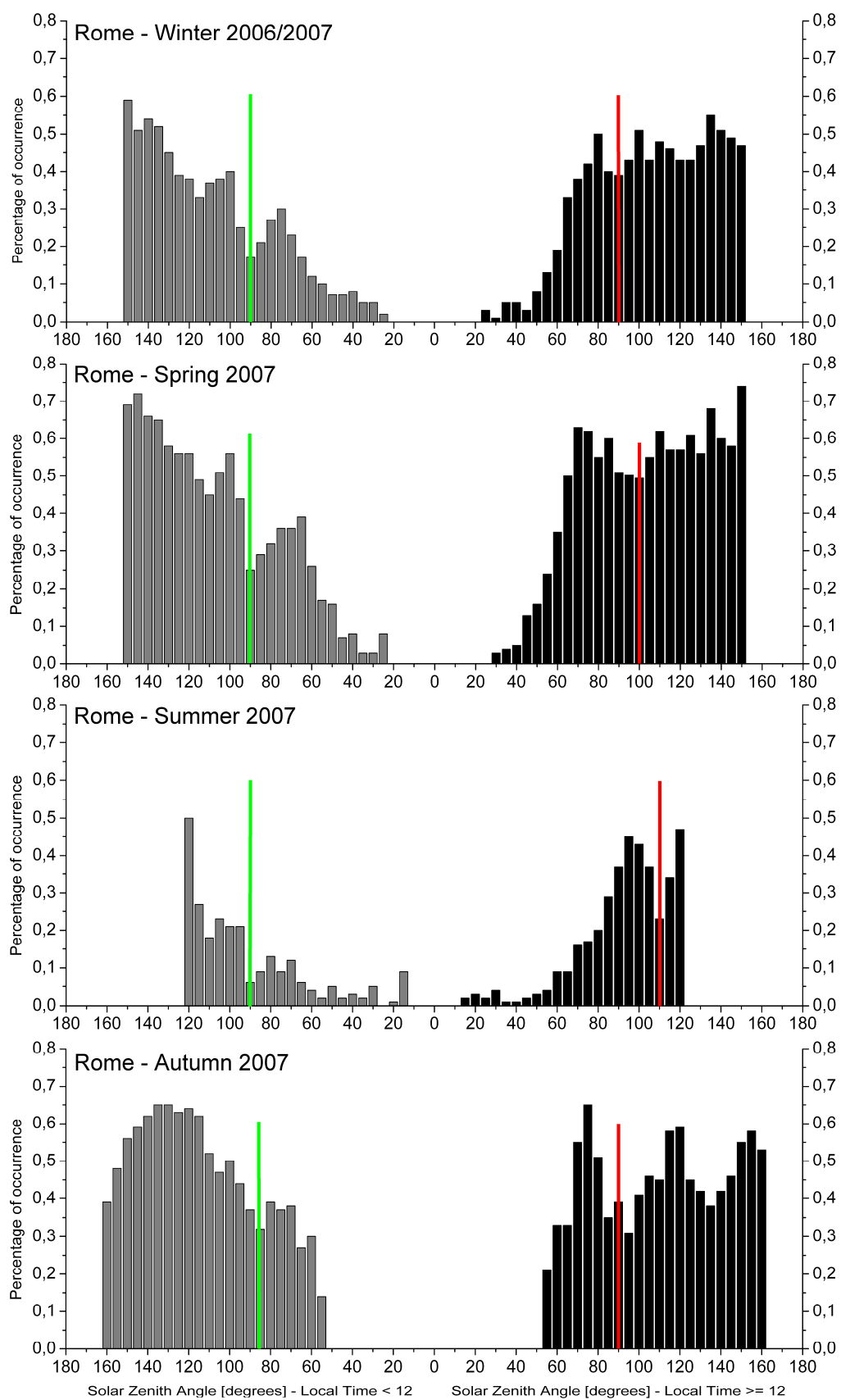


Figure 4

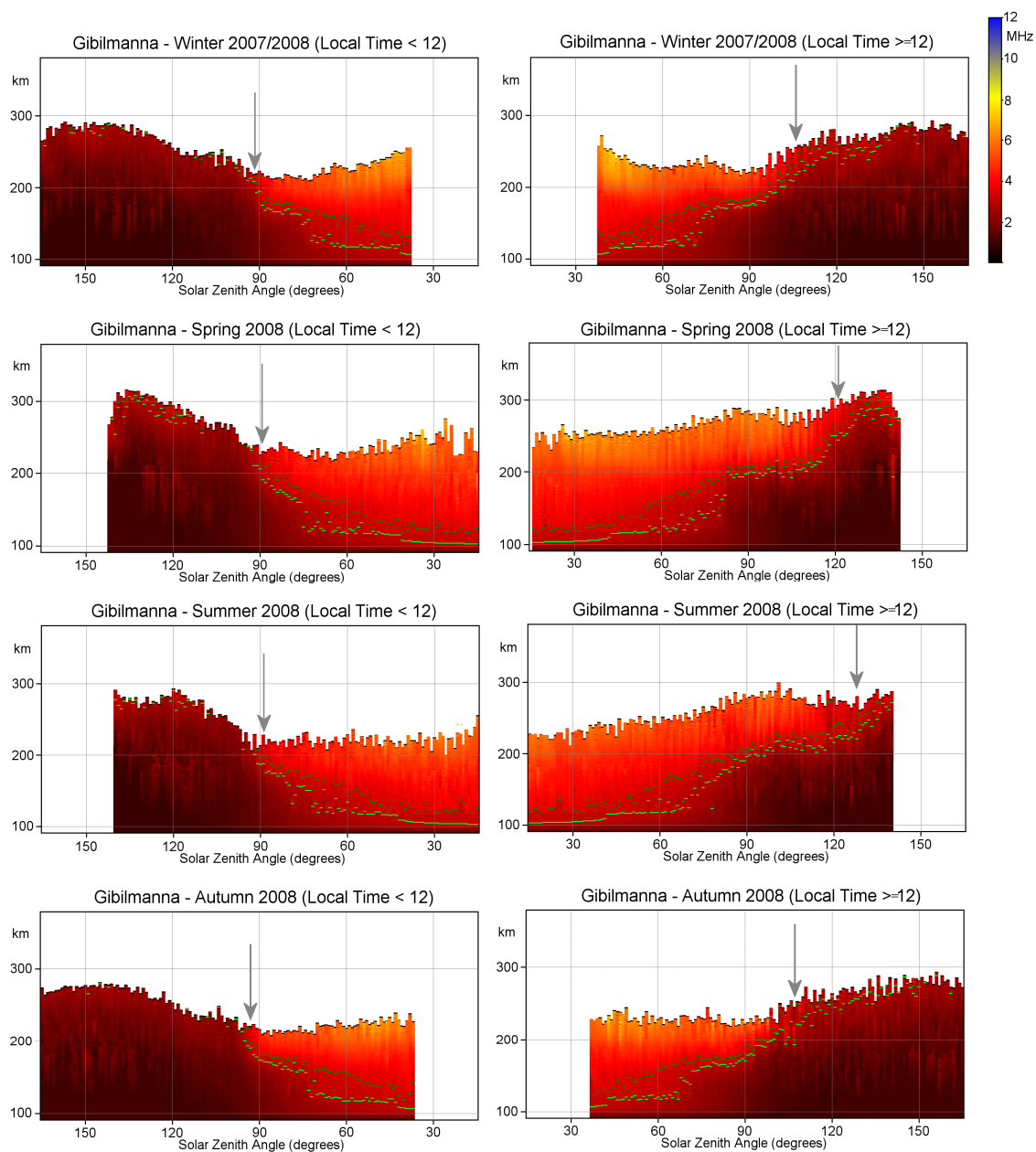


Figure 5

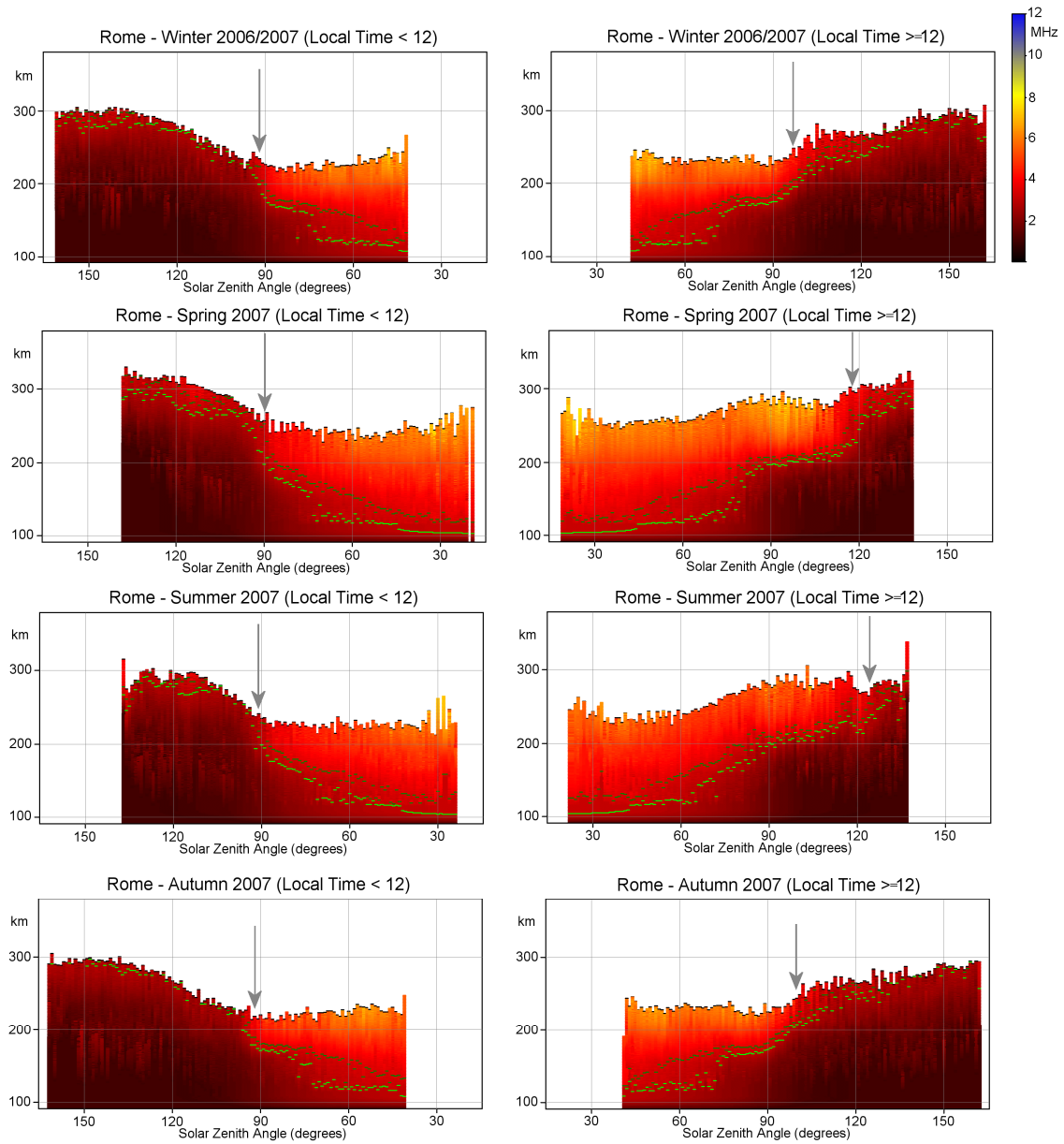


Figure 6